DOCUMENT RESUME

ED 144 988 TM 006 551

AUTHOR Duszynska, Anna
TITLE Reasoning Level Test. Application of Piaget's Theoretical Model to the Construction of a Science Test for Elementary School.
NOTE 28p.

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS Academic Achievement; Age Differences; *Cognitive Development; Comparative Analysis; *Cross Cultural Studies; Elementary Education; Elementary School Students; *Logical Thinking; Multiple Choice Tests; Racial Differences; Reading Comprehension; School Location; Sciences; *Science Tests; Scientific Literacy; Sex Differences; *Test Construction
IDENTIFIERS *Piagetian Theory; Poland; *Science Reasoning Level Test (Duszynska); United States

ABSTRACT

The Science Reasoning Level Test (SRLT) was developed and administered to elementary school students in Poland and the United States, in order to determine whether a multiple choice test based on a Piaget model of the development of thought processes could be used to study the development of scientific reasoning. The relationship between reasoning and academic achievement, and patterns in the development of thought processes were compared for both countries. Test scores for 171 Polish children in grades 3-5 and 70 American children in grades 4-6 revealed positive correlations between the SRLT and scores on science and reading comprehension tests in both countries. A positive correlation was also found between Guttman scales scores, used to evaluate patterns of responses, and school achievement. There were no significant differences in performance between Polish and American students. It was demonstrated that Piaget's conceptualization of stages of thinking can be used as a basis for test construction. Comments on the instructional applications of such tests, and a copy of the Science Reasoning Level Test are included. (Author/MV)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED144988

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

REASONING LEVEL TEST

Application of Piaget's theoretical model to the construction of a science test for elementary school.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Anna Duszynska

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

DUSZYNSKA, ANNA

INSTITUTE OF EDUCATIONAL RESEARCH, UL GORCZEWSKA8

01-180 WARSAW, POLAND

TM006 551

ABSTRACT

The main objective of this study was the application of Piaget theory to educational research. Piaget's model of the development of logical thinking was used here as a basis for construction of the Science Reasoning Level Test. The responses in multiple choice items were classified into stages of thinking: either preoperational, concrete, or formal. 170 Polish elementary school students were tested in Poland and 70 students from elementary schools in the United States. The Guttman scalogram analysis was applied to evaluate the patterns of responses. For both Polish and American students reproducible scales were obtained. The correlation between the level of scientific reasoning and school achievement in science and reading comprehension was studied. The results show positive correlations between Science Reasoning Test, Science, and Reading Comprehension Test scores in both countries. Positive correlation has also been found between Guttman scales scores and school achievement. No significant differences exist between the performance of Polish and American students. It was demonstrated that Piaget model is a good basis for test construction. Testing gives the teacher the opportunity to detect patterns of thinking, to classify children according to their choice of response into certain stages of thinking and finally to introduce different modes of instruction according to the information provided by the results of testing.

INTRODUCTION

For several years we have observed a constant increase in interest in psychological aspects of learning and teaching laboratory sciences. This happened because of the improvement in instruction of general cognitive skills, especially after laboratory activities were introduced into science curricula. The general belief of curriculum reformers was that experience helps students to develop cognitive ability that can ultimately lead them to think abstractly. But in spite of possible advantages of this method of instruction, students still get the lowest scores when solving the items in which scientific hypothetical deductive reasoning is required.[1] Studies show that only 50% of high school and college students have aquired formal reasoning skills while studying science. It seems clear that the analysis of the process of learning science, especially through practical laboratory skills should be carried out. Such an analysis could provide essential information about the development of mental operations involved. Diagnosis of learning characteristics in terms of the level of reasoning ability is necessary prior to his entry into laboratory activities. For the purposes of analyzing the process of reasoning a Piaget model can be applied. It has been demonstrated that the logical model of Piaget can be used as a basis for paper and pencil test construction.[2] Such testing reveals the occurrence of a fixed sequence of stages in the development of logical thinking in children. Similar results were obtained in countries differing greatly in language and culture.[2]

PROBLEM

In the present study the following problems were taken into consideration:

1. Can a multiple choice test based on a Piaget model be used to study the development of scientific reasoning in children?
2. Is the level of scientific reasoning positively correlated with school achievement?
3. Does the level of reasoning follow the same pattern in two different countries such as Poland and the United States?

In order to answer these questions the Science Reasoning Level Test was constructed and administered in three elementary schools in Poland and one elementary school in the United States.

DESCRIPTION OF THE TEST

Originally a 25-item questionnaire was administered to some 50 Polish children. Most of the items represented classical Piagetian tasks that have been developed and used by Piaget and his co-workers.[3] Other items were constructed on the basis of Piaget's model and they also meet the requirement of hypothetical deductive reasoning. All items were presented as open-ended questions. The main source of the alternative choices in multiple-choice items of the final test were the responses children reproduced themselves while answering open-ended

questions in the pilot study. The children's responses were classified into stages of logical thinking according to Piagetian theory. Of the 25 items, 14 were used for the construction of Science Reasoning Level Test. All the items in the Science Reasoning Level test are based on simple experimental situations* and encourage students to reason about concrete biological or laboratory problems.

The Sample.

The group used for testing in Poland consisted of 171 elementary school children from the 3rd, 4th and 5th grade. The age range was 9-12. In the American school the Science Reasoning Level Test was administered in 4th, 5th and 6th grades to 70 children. 24 children were black and 46 were white. The age range was 9-12. One measure of SRLT was used in present study-total weighted score.

PROCEDURES AND RESULTS

At the beginning an item analysis was carried out. First to show that the total weighted score varies, as we would expect, with responses on particular items and, second, to provide basic item - total test information.

All children were classified, separately for every item, into three groups, either preoperational, concrete, or formal according to their choice of response. The mean scores and standard deviations for groups are presented in Table I. The finding that the differences

*finding key variable, discovering the idea of controlled experiment, discovering the cause and effect relationship, correlation, relating in simple ecosystems, etc.

among group means were significant shows that there is a development in expected direction from preoperational to formal reasoning.

This supposition was next tested by means of Guttman scalogram analysis. The hypotheses to be tested were:

1. Preoperational thinking would precede two higher stages of thinking.

2. Preoperational and concrete thinking would procede the highest formal thinking. For the purposes of scalogram analysis the items of the SRLT were changed into dichotomies, by putting different cut-off points. Two different cut-off points were tried.

   a) cut-off point "formal" -- divides the responses into two categories, giving the score of 1 (pass) to the responses based on formal thinking and score of 0 (fail) to the choices based on either concrete or preoperational thinking.

   b) cut-off point "concrete" -- gives the score of 0 to preoperational choices, and the score of 1 to concrete and formal responses.

Each cut-off point gives two different values of difficulty (% of correct responses) to each item.

Therefore, items were layed out twice in each grade according to their difficulty, and each time two scales were constructed - "formal" and "concrete." Subjects were ranked according to their scale scores calculated from relevant scales. In Table 2 coefficients of reproducibility, together with measures of reliability in the Polish sample are

presented. The scale was accepted as successful if it gave a coefficient of reproducibility not lower than 0.85. All the scales possess the required value of the coefficient. The results support both hypotheses.

If a cross cultural uniform pattern of responding to the test occurs, then the same Guttman scales will function in Poland and the United States. Six new scales were constructed for American students with the order of difficulty transferred from the relevant Polish scales. Talbe 3 shows the coefficients of reproducibility of the 6 American scales.

By comparison of individual scale scores on "formal" and "concrete" scales children can be classified into 4 possible groups.

1. Children low on formal and low on concrete scale. (preoperational)
2. Children high on both scales. (formal)
3. Children low on formal and high on concrete scale. (concrete)
4. Children high on formal and low on concrete. (different thinking)

The scale score was considered low if it was located on the lower half of the scale, and high if it was located on the upper half of the scale. It appears that the majority of the students in all three grades fall into either preoperational or formal stage (about 60 to 80%). Specific concrete thinkers appear in very small number.

The existence of scales proves to a certain extent that there exist types of thinkers, that students develop certain reasoning patterns and apply them to different tasks. But those results prove that such patterns would probably differ greatly depending on the task presented to the children. Nevertheless, both the results of scalogram analysis and item analysis

demonstrate that multiple choice items can be successfully used as a measure of development of scientific reasoning in children.

Next the relationship between the level of reasoning and school achievement was studied. Polish students were additionally tested at the time of main testing on their school achievement in science and reading comprehension. Reading Comprehension Test and Science Test (both constructed for the purpose of the international survey conducted by International Association for the Evaluation of Educational Achievement) were chosen as external criteria. For the American sample the IQ scores were used for correlation study. In Table 4 correlation coefficients are presented between science, reading, comprehension and scientific reasoning in the Polish sample. In Table 5 correlation coefficients between IQ and scientific reasoning are presented for the American sample. Also scale scores are positively correlated with measures of school achievement (Table 6). It seems that the relationship between reasoning ability and school achievement increases through grades. With years spent in school, logical thinking becomes more and more related to school achievement. The correlation between reading comprehension and science remains the same through grades.

The results of the correlation analysis support the fact that the test and scales are reliable* and that they measure what they were designed for, i.e. the process of scientific reasoning. If the coefficients were very high, the test would probably be another measure of school achievement, if the coefficients were low it would prove that the test was measuring a process totally independent of school performance.

*The correlation coefficients can be regarded as lower bounds to the reliability.

The SRLT and Guttman scales were tested for their test-retest reliability (Table 7). The coefficients of reliability are .66 for the Polish sample and .51 for the American. With respect to reliability both the split-self lower bound and the standard error of reproducibility were completed. Unreliability is but one possible reason why variables are uncorrelated. They may be uncorrelated even though they are scalable. Scalogram analysis provides as an automatic by-product the assurance that the responses to individual items have relatively little error of measurement if reproducibility is high.

The level of thinking seems to be dependent on sex, kind of school, and age. In the Polish sample, children from urban schools, mostly coming from middle class homes, show better performance than the students from suburban and rural schools. Boys performed better than girls. The effects of school, grades and sex have significant effect on performance in the Science Reasoning Level Test. In the American sample no significant differences were found between sexes or between black and white children. A significant difference was found only between the 4th and 6th grades. The Polish and American data were compared with regard to total weighted score and scale scores. No significant differences were found through grades and sexes (Table 7, 8).

CONCLUSIONS

It was demonstrated that Piaget's model can be applied to the measurement of reasoning involved in learning laboratory sciences. Piaget theory can be used as a basis for test construction. By test examination the occurrence of stages can be revealed even in countries differing greatly

in language and culture. The application of Piaget theory to test measurement has important advantages for teachers and test constructors.

First, before construction of alternative choices, the children themselves should be studied as to their modes of thinking and their mistakes. When constructing multiple-choice items the teacher does not look for responses through random guesses as to possible mistakes, but has clear evidence of them.

Second, the Piaget logical model gives a theoretical basis for classification of children's responses with regard to the mental operations involved.

Third, the results of testing give him an opportunity to detect different patterns of thinking, to classify children according to their choice of responses into different stages of thinking or to groups of particular thinkers. This in turn helps to introduce relevant methods of instruction, especially during laboratory activities.

It is very imporant at early stages of education to analyze the cognitive skills of students, in terms of the level of reasoning ability (formulating hypotheses, understanding the meaning of controlled experiment, of cause-effect relationships, etc.) prior to his entry into laboratory activities. Measuring global notions of intelligence is not useful because it obscures specific individual differences.

The results of such testing can help to reveal why in spite of the general belief that experience in the laboratory helps students to develop cognitive skills, they still get lower scores when scientific reasoning is required.

REFERENCES

1. Comber, L. C. and Keeves, J. P. Science Education in Nineteen Countries: An Imperical Study, International Studies in Evaluation, Stockholm, 1973, Almquist and Wiksell International.
2. Bergling, K. The Development of Hypothetical-Deductive Thinking in Children, Stockholm, 1974, Almquist and Wiksell International.
3. Inhelder, B., and Piaget, T., The growth of logical thinking from childhood to adolescence, London, RouHege and Kegan, Paul, Ltd., 1958.
4. Stouffer, S.A. et al (eds) Measurement and Prediction.
5. Princeton, Princeton University Press, 1950.

TABLE 1

Total weighted score by response (preoperational, concrete or formal) within items of the Science Reasoning Level Test. Polish sample.

	(N) Stage			Mean Weighted Score			Standard Deviation		
Item	1	2	3	1	2	3	1	2	3
1	41	32	96	1.91	2.05	2.21	0.29	0.28	0.28
3	42	97	31	1.91	2.15	2.21	0.30	0.28	0.30
4	21	85	58	1.94	2.04	2.27	0.33	0.28	0.27
5	82	35	54	1.98	2.21	2.22	0.28	0.32	0.28
6	42	23	105	1.84	2.01	2.23	0.29	0.20	0.27
8	75	41	55	1.94	2.17	2.28	0.29	0.26	0.26
9*	44	42	36	1.96	2.15	2.27	0.25	0.28	0.29
10	22	43	105	1.93	2.04	2.17	0.27	0.35	0.28
11	59	31	81	2.01	2.02	2.20	0.30	0.32	0.29
12	37	49	80	1.86	2.04	2.25	0.29	0.24	0.27
13	28	63	75	1.88	1.98	2.29	0.25	0.24	0.27
14	49	40	81	1.88	2.10	2.24	0.25	0.28	0.28
15	68	71	31	2.09	2.08	2.20	0.33	0.30	0.29
16	51	96	26	1.99	2.14	2.20	0.30	0.31	0.24

1 - preoperational stage

2 - concrete stage

3 - formal stage

*Children in 4th grade didn't attempt item 9

TABLE 2

Coefficients of reproducibility, standard errors of the coefficients and reliability coefficients (split-half lower bound) for 6 Guttman scales. Polish sample.

Grade	Scale	Coefficient or reproducibility	Standard error	Split-half lower bound
3	formal	.87	.031	.46
	concrete	.87	.031	.60
4	formal	.86	.044	.54
	concrete	.83	.044	.44
5	formal	.86	.031	.60
	concrete	.87	.031	.40

TABLE 3

Coefficients of reproducibility and standard errors of reproducibility for Guttman scales. American sample.

Grade	Scale	Coefficient of reproducibility	Standard error
4	formal	.80	.078
	concrete	.78	.081
5	formal	.81	.070
	concrete	.87	.065
6	formal	.81	.092
	concrete	.85	.084

TABLE 4

Pearson correlation coefficients and multiple correlation coefficients between the Science Reasoning Level Test and Reading Comprehension and between Science Reasoning Level Test and Science Test.

	GRADE		
Criterion (SRLT)	3	4	5
Science	.35	.52	.58
Reading Comprehension	.32	.63	.51
Multiple r	.38	.64	.62

TABLE 5

Pearson correlation coefficients for IQ, total weighted score and scale scores. American sample.

	IQ
Total weighted Score	.40
Scale formal	.43
Scale concrete	.35

TABLE 6

Pearson correlation coefficients and multiple correlation coefficients between the Guttman scales and the Reading Comprehension and Science tests.

	GRADE					
	3		4		5	
Test	scale formal	scale concrete	scale formal	scale concrete	scale formal	scale concrete
Reading Comprehension	.54	.26	.69	.27	.51	.38
Science	.45	.31	.52	.51	.56	.48
Multiple r	.55	.29	.69	.54	.61	.50

TABLE 7

Means and standard deviations by grades. Poland and USA.

	GRADE								
	3 (4)			4 (5)			5 (6)		
	X	SD	N	X	SD	N	X	SD	N
USA	1.97	0.26	26	2.17	0.10	26	2.19	0.22	18
POLAND	2.01	0.30	59	2.09	0.31	49	2.19	0.29	63

TABLE 9

Means and standard deviations for boys and girls. Poland and USA.

	Girls			Boys			White girls			White boys		
	X	SD	N	X	SD	N	X	SD	N	X	SD	N
USA	2.04	0.30	83	2.17	0.31	81						
POLAND	2.05	0.25	31	2.14	0.29	39	2.09	0.24	20	2.18	0.26	26

REASONING LEVEL TEST

DIRECTIONS

Circle the correct answer in each part.

Answer every question, even if you are not sure.

If you have to skip go back later.

Please, don't guess.

NAME: _____

GRADE: _____

AGE: _____

* P - Preoperational thinking
** F - Formal operational thinking
*** C - Concrete operational thinking

1. Some seeds sprout better in the dark, others in the light, and some sprout just as well in the dark as in the light. Joe wanted to check which kind of seeds he had. What should he do?

   P.* Sow some seeds on the wet blotting paper and keep them in the dark.

   P. Sow some seeds on the wet blotting paper and keep them in the light.

   F.***Sow some seeds on the wet blotting paper, keep one part in the dark and the other part in the light.

   C.** Sow some seeds on the wet blotting paper, keep them in the dark for awhile, and then put them in the light.

3. Once upon a time there lived deer, lions and wolves at the banks of a river. Hunters killed all lions and wolves in that area. What would happen to the deer?

   C. The deer would devastate all plants at the river.

   P. People killed lions and wolves because these are very dangerous animals.

   C. More and more deer began to live at the banks of that river.

   F. Many deer died of starvation.

4. A mule and a donkey have long ears. A horse and a mule have long tails. Which animal has long ears and long tail?

   C. A donkey or a mule.

   C. A horse or a mule.

   P. A dog.

   F. A mule.

5. Some children collected a bunch of 50 flowers from a garden. The bunch consisted of 20 red and sweet smelling flowers, 1 white and sweet smelling flower, two red and nonsweet smelling flowers and 27 white and nonsweet smelling flowers. Ann said, "The color of flowers determines if they smell or not, because the flowers are mostly either red and sweet smelling or white and nonsweet smelling."
Then Mike said, "There is no connection between color and smell. If the flowers are well watered they smell sweet."
But Mary said, "There is no connection between the color of flowers and their smell, because there are two red nonfragrant flowers and one white frangrant flower in the bunch."
Then Bill said, "There are many more red, sweet smelling flowers than red nonsmelling ones, so there is a connection between color and smell."
Who was right?

P. Mike.

C. Mary.

F. Ann.

F. Bill.

6. There were two pigeon-coops. The pigeons from one pigeon-coop were fed with an unmilled rice and water only. The pigeons from the second coop were fed with milled rice and water only. You probably know that there is a nutritious substance in unmilled rice, meat, nuts and wheat which is necessary for life. The pigeons from the

first coop grew well and were healthy. The pigeons in the second coop ate up all their food but didn't grow well and became ill. Why were the pigeons in the second coop sick?

P. Because of bacteria.

F. Because they lacked that nutritious substance.

C. Milled rice was bad for them.

P. Strong pigeons forced the small pigeons from the food and they became ill.

8. Once upon a time, on a small island, there lived only cats, birds, ladybugs and little beetles. The beetles ate plants only, ladybugs lived on beetles only, birds ate only ladybugs, and cats ate only birds. What would happen if all the ladybugs disappeared from that island?

C. Birds would be hungry.

P. There would be only beetles, birds, and cats on that island.

F. Cats would starve to death and all plants would be destroyed by beetles.

P. There would be no ladybugs, which are beautiful and useful insects.

9. There were four cages with rabbits in the laboratory. The rabbits from the first cage were getting food without fat, without salt but with sugar. The rabbits from the second cage were getting food without sugar, without fat but with salt. The rabbits from the

third cage were getting their food without sugar but with fat and with salt. The rabbits from the fourth cage lived on food without sugar, without salt but with fat.

Why did the rabbits in the cages number 1 and 4 become ill?

P. One rabbit was ill and infected other rabbits.

P. The rabbits became ill because they didn't want to eat food without sugar or salt, it tasted badly.

C. The rabbits became ill because they didn't get salt and fat.

F. The rabbits became ill because they didn't eat enough salt.

10. Kate grew four flowers: blue, red, yellow and white in the four flowerpots. She kept the blue and red flowers in the dark. She watered the blue flower but she didn't water the red one. Kate put the yellow and the white flowers on the windowsill. She watered the yellow one but she didn't water the white one.

Which flowers began to wilt and why?

C. Blue and red because they lacked light.

P. The flowers began to wilt because she used poor soil.

F. Blue, red and white flowers wilted because they didn't get enough light and water.

C. The white flower wilted because it lacked water and the blue because it lacked light.

11. People and animals eat, digest and break down their food in their bodies to obtain energy necessary for life. The elephant weighs 2000 pounds and breaks down his food with a speed equal to 100 grams per

hour. A mouse weighs 50 grams and breaks down his food with a speed equal to 1500 grams per hour. A dog weighs 20 pounds and does it with a speed equal to 400 grams per hour. A grown man weighs 160 pounds and breaks down his food with a speed equal to 200 grams per hour.

What does that speed depend on?

P. The speed depends on the kind of food the animal lives on.

C. An elephant weighs more than a man. A man weighs more than a dog, a dog weighs more than a mouse. A mouse breaks his food more quickly than a dog and a dog more quickly than a man. A man does it more quickly than an elephant.

F. The more an animal eats the more quickly it breaks down its food.

P. The less the man or the animal weighs the more quickly he breaks down his food.

12. Marigolds need more light to grow than ivy, but less than roses. Which of these plants needs the least light of all, and which the most?

C. Marigold the most, ivy the least.

F. Ivy the least, roses the most.

P. Sunflower.

C. Roses the least, ivy the most.

13. Sharks and whales live in the ocean. Llamas and whales breath with their lungs. Which animal lives in the ocean and breathes

with his lungs?

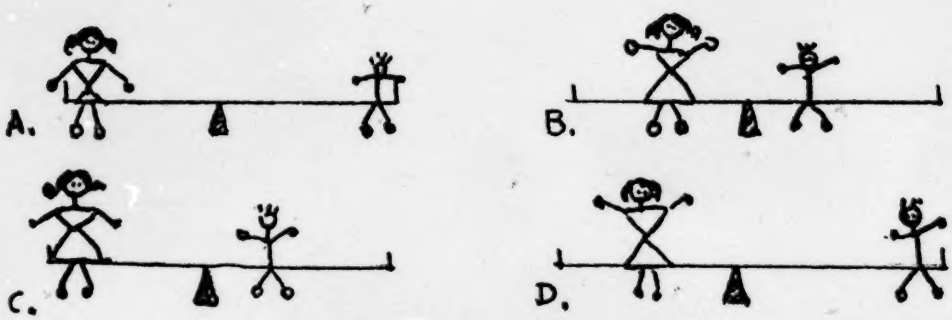
C. A whale or a shark.

P. A fish.

F. A whale.

C. A whale or a llama.

14. Betty wanted to play on a see-saw with her little brother John. How should they be seated to maintain the balance on a see-saw?



15. There were yellow and green beans in a paper bag. One type of beans had smooth skin and the other wrinkled. There were 14 smooth and yellow beans, 16 green and wrinkled, 15 green and smooth and also 15 yellow and wrinkled.

Is there any connection between the color and texture of beans?

P. No, the color and texture depend on the type of soil used to grow the beans.

C. Yes, there is a connection between the color of beans and the texture, because there are almost as many yellow and smooth beans as green and wrinkled, and as many green and smooth as yellow and wrinkled.

F. There is no connection between the color of beans and their texture because there are as many yellow and smooth beans as yellow

and wrinkled, and as many green and smooth beans as green and wrinkled.

C. There is a connection because half or the beans are yellow, and half of them are wrinkled.

16. Tom wanted to know which of three typical kinds of soil is the best for raising beans: normal, clay, or sandy soil.
He found three flower pots one was big, one was middle size and one was small. He filled each pot with a different kind of soil and grew the same number of beans in each of the pots. Then he put the pots one by one on a windowsill and watered them with equal amounts of water.
Why was Tom's experiment bad?

P. None of those soils are good for raising plants.

F. There was a different amount of soil in each flower pot.

P. Sandy soil needs more water than clay soil.

C. Tom should have poured more water into the big pot than into the small one.